

UNITED STATES PATENT APPLICATION

FOR

Method for Secure Key Exchange

INVENTOR:

Balaji Vembu

INTEL CORPORATION

Prepared by:

Steven P. Skabrat

Reg. No. 36,279

(503) 264-8074

Express Mail No EV325528715US

Method for Secure Key Exchange

5 BACKGROUND

1. FIELD

The present invention relates generally to computer security and, more specifically, to exchanging cryptographic keys in a processing system.

10

2. DESCRIPTION

One of the hurdles in providing protected digital content on a computing platform (such as the personal computer (PC)) is that the application program that is extracting the protected content and the graphics device that is decoding 15 and/or displaying the content need to agree on a cryptographic key to encrypt the data exchange between them. If the content is not encrypted during transfer between the application and the graphics device, the content may be vulnerable to interception. One of the two entities cannot merely generate the key and send the key to the other entity because there is typically no “non- 20 snoopable” secure path between the application and the graphics device.

One approach to this problem is to embed identical encryption keys in the graphics device and the application, and then use this key. This approach entirely avoids the key exchange. However, this solution is not robust because the application program may be hacked to discover the key. Another approach 25 is to embed a private key of a public/private key pair in the graphics device and send the corresponding public key to the application. The application then uses the public key to encrypt the content and the graphics device uses the private key to decrypt the content. Alternatively, the application can generate a new symmetric session key, encrypt it with the graphic device's public key, and send 30 it to the graphics device (hence the public-private key pair is used to enable exchange of the symmetric key). Since by definition the public key does not

need to be protected from other agents, the public key can be sent to the application in the clear.

Both of these approaches require a key to be embedded in the graphics device. This increases the manufacturing cost and adds complexity to the manufacturing flow for graphics devices. A better approach is needed.

BRIEF DESCRIPTION OF THE DRAWINGS

10 The features and advantages of the present invention will become apparent from the following detailed description of the present invention in which:

Figure 1 is a diagram illustrating a processing system having a trusted platform module (TPM) according to an embodiment of the present invention;

15 Figure 2 is a flow diagram illustrating a process for secure key exchange according to an embodiment of the present invention; and

Figure 3 is a diagram illustrating communications between an application, a graphics device and a TPM according to an embodiment of the present invention.

20

DETAILED DESCRIPTION

An embodiment of the present invention is a method of exchanging a 25 cryptographic key between two entities within a processing system. In one embodiment, the two entities may be an application program and a graphics device. However, embodiments of the present invention may be used with any two entities in a processing system. Embodiments of the present invention make use of a trusted platform module (TPM), which is used as a root of trust for a 30 processing system. Every TPM contains a public/private key pair. Embodiments of the present invention leverage the TPM's key pair to facilitate the key

exchange. The present invention does not require any keys to be embedded in the entities that need to agree on a common key.

Reference in the specification to "one embodiment" or "an embodiment" of the present invention means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, the appearances of the phrase "in one embodiment" appearing in various places throughout the specification are not necessarily all referring to the same embodiment.

An exemplary processing system for embodiments of the present invention is shown in Figure 1, however, other systems may also be used and not all components of the processing system shown are required for the present invention. Sample system 100 may be used, for example, to execute the processing for embodiments of the present invention. Sample system 100 is representative of processing systems based on the PENTIUM® family of processors and CELERON™ processors available from Intel Corporation, although other systems (including personal computers (PCs) or servers having other processors, engineering workstations, other set-top boxes, and the like) and architectures may also be used.

Figure 1 is a block diagram of a system 100 of one embodiment of the present invention. The system 100 includes a processor 102 that processes data signals. Processor 102 may be coupled to a processor bus 104 that transmits data signals between processor 102 and other components in the system 100. System 100 includes a memory 106. Memory 106 may store instructions and/or data represented by data signals that may be executed by processor 102. The instructions and/or data may comprise code for performing any and/or all of the techniques of the present invention. Memory 106 may also contain additional software and/or data such as at least one application program 107.

A bridge/memory controller 110 may be coupled to the processor bus 104 and memory 106. The bridge/memory controller 110 directs data signals between processor 102, memory 106, and other components in the system 100

and bridges the data signals between processor bus 104, memory 106, and a first input/output (I/O) bus 112. In this embodiment, graphics device 113 interfaces to a display device (not shown) for displaying images rendered or otherwise processed by the graphics device 113 to a user. Graphics device 113

5 may comprise a start key exchange register 115. The graphics device may receive data such as protected digital content from the application when the application is being executed by the processor.

First I/O bus 112 may comprise a single bus or a combination of multiple buses. First I/O bus 112 provides communication links between components in 10 system 100.

In at least one embodiment, a trusted platform module (TPM) 116 may be coupled to bus bridge 126. A TPM comprises circuitry included within a processing system to support trusted computing. A TPM has been defined by the Trusted Computing Group (TCG) in the Trusted Computing Platform

15 Association (TCPA) Main Specification 1.2, February 2002, and successive versions, available from the TCG. A TPM operates somewhat like a "smart card" on a motherboard of a computer system (such as a personal computer (PC)), to provide various security functions to the system. There is only one

TPM per system. The TPM includes at least one public/private key pair for

20 use in cryptographic operations, can generate anonymous key pairs for use by other entities within the system, can perform encryption and decryption operations, can sign and verify data, and can establish a root of trust for the system. The TPM is considered to be difficult to break into and affect its operations. In at least one embodiment, TPM 116 comprises at least three

25 registers, denoted register A 117, register B 118, and register C 119, herein. Embodiments of the present invention use the TPM's key infrastructure and these three registers to perform a key exchange.

A second I/O bus 120 may comprise a single bus or a combination of multiple buses. The second I/O bus 120 provides communication links between

30 components in system 100. A data storage device 122 may be coupled to the

second I/O bus 120. A keyboard interface 124 may be coupled to the second I/O bus 120. A user input interface 125 may be coupled to the second I/O bus 120. The user input interface may be coupled to a user input device, such as a remote control, mouse, joystick, or trackball, for example, to provide input data to 5 the system. A bus bridge 126 couples first I/O bridge 112 to second I/O bridge 120.

Embodiments of the present invention are related to the use of the system 100 as a component in a content protection and rendering system. According to one embodiment, such processing may be performed by the system 100 in 10 response to processor 102 executing sequences of instructions in memory 106. Such instructions may be read into memory 106 from another computer-readable medium, such as data storage device 122, for example. Execution of the sequences of instructions causes processor 102 to execute secure key exchange processing for the application according to embodiments of the 15 present invention. In an alternative embodiment, hardware circuitry may be used in place of or in combination with software instructions to implement portions of embodiments of the present invention. Thus, the present invention is not limited to any specific combination of hardware circuitry and software.

The elements of system 100 perform their conventional functions in a 20 manner well-known in the art. In particular, data storage device 122 may be used to provide long-term storage for the executable instructions and data structures for embodiments of components of the content distribution system in accordance with the present invention, whereas memory 106 is used to store on a shorter term basis the executable instructions of embodiments of components 25 of the content distribution system in accordance with the present invention during execution by processor 102.

Figure 2 is a flow diagram illustrating a process for secure key exchange between entities in a processing system according to an embodiment of the present invention. At block 200, a first entity in the processing system, such as 30 application program 107 for example (which in some embodiments may comprise a content player application) signals a second entity, such as graphics

device 113, to start key exchange processing. The key exchange may be needed to support protected transmission of content from the first entity to the second entity. In one embodiment, the second entity comprises a hardware device (such as a graphics controller/graphics card/graphics device) having a 5 register that, when written to, indicates a request to start key exchange processing between the hardware device and another entity in the processing system. In one embodiment, this register may be known as a start key exchange register 115, and the first entity (e.g., the application) causes the writing of a predetermined value into the start key exchange register to start key exchange 10 processing. The writing of the start key exchange register is represented as flow 300 on Figure 3. In other embodiments, other methods of signaling the graphics device to start the key exchange may be used.

At block 202, after detecting the writing of the start key exchange register, the graphics device generates a first pseudorandom symmetric key K_D using any 15 well known method of key generation. The TPM comprises at least one public/private key pair. The TPM's public key is known by other entities in the processing system. At block 204, the graphics device encrypts the pseudorandomly generated key K_D using the TPM's public key $K_{TPM-PUB}$ according to well known methods of public key cryptography to form a first 20 encrypted value $E(K_D, K_{TPM-PUB})$. At block 206, the graphics device sends the first encrypted value $E(K_D, K_{TPM-PUB})$ to TPM 116 and causes the writing of the first encrypted value $E(K_D, K_{TPM-PUB})$ into a first register, such as Register A 117. This action is represented as flow 302 on Figure 3. Since the TPM's public key 25 was used to encrypt the graphic device's key, only the TPM's corresponding private key can be used to decrypt the graphic device's key.

Next, at block 208, the application generates a second pseudorandom symmetric key K_A using any well known method of key generation. At block 210, the application encrypts the pseudorandomly generated key K_A using the TPM's public key $K_{TPM-PUB}$ according to well known methods of public key cryptography 30 to form a second encrypted value $E(K_A, K_{TPM-PUB})$. At block 212, the application sends the second encrypted value $E(K_A, K_{TPM-PUB})$ to TPM 116 and causes the

writing of the second encrypted value $E(K_A, K_{TPM-PUB})$ into a second register, such as Register B 118. This action is represented as flow 304 on Figure 3. Since the TPM's public key was used to encrypt the application's key, only the TPM's corresponding private key can be used to decrypt the application's key.

5 In one embodiment, blocks 208 to 212 may be performed concurrently with blocks 200 to 206.

At block 214, the TPM decrypts the first and second encrypted values stored in register A and register B, respectively, using the TPM's private key $K_{TPM-PRI}$. The TPM now has both the application's key K_A and the graphic

10 device's key K_D . Since the TPM is very difficult to access in an unauthorized way, these values are considered to be secure. Next, at block 216, the TPM encrypts the graphic device's symmetric key K_D using the application's symmetric key K_A and an appropriate encryption algorithm, and stores the third encrypted value $E(K_D, K_A)$ into a third register in the TPM, such as register C 119. At block 15 216, the application reads the contents of register C 119 from the TPM to obtain the third encrypted value $E(K_D, K_A)$. This action is represented as flow 306 in Figure 3. The capability for the application to obtain the contents of register C from the TPM is assumed in this embodiment. At block 218, the application decrypts the contents of register C received from the TPM using the application's 20 own key K_A and a corresponding decryption algorithm to obtain K_D .

Since all of the data flows between entities as shown in Figure 3 were encrypted, the data may be secure against hackers attempting to discover the keys by monitoring communications between the entities.

25 The graphic device's symmetric key K_D may now be used by the application to encrypt content. The encrypted content may be sent by the application to the graphics device, where the encrypted content may be decrypted by the graphics device and further processed (e.g., rendered for perception by the user), since the graphics device already knows its own key K_D .

30 In one embodiment, if the graphics device needed to know the application's key K_A , a similar mechanism to blocks 216 to 220 may be used to transfer the application's key to the graphics device in a protected manner.

In the above embodiments, for good security the application needs to ensure that the key exchange happened with a legitimate graphics device and not a hacker's unauthorized device masquerading as the graphics device. This can be ensured by providing a private path from the graphics device to a 5 dedicated TPM graphics processor input/output (I/O) (GPIO) pin. This path would ensure that Register A 117 would be written only with data received over the line connected to the GPIO pin to the TPM. The GPIO pin may be connected to the graphics device (e.g., an advanced graphics processor (AGP)/peripheral component interconnect (PCI) slot) using a buried line on the 10 printed circuit board of the processing system's motherboard that is not easily detectable by a hacker. This line is also represented as flow 302 on Figure 3.

In the absence of embedded keys in the endpoints of the exchange, if two endpoints want to agree on a common key it typically requires a fairly elaborate and complex protocol (e.g., Diffie-Hellman key exchange). One simple solution 15 is to embed keys in the endpoints and then use the embedded key to encrypt the key that needs to be exchanged. However, embedding keys in the endpoints adds a lot of security overhead because the keys have to be protected. Additionally, embedded keys raises revocation issues for devices. If the keys need to be unique, then this adds to the manufacturing complexity and cost. In 20 contrast, embodiments of the present invention allow a secure key exchange to take place between two entities (e.g., an application program and a graphics device) without either of them embedding any keys. The entities take advantage of the public/private key pair that is available on the platform in the TPM to perform the key exchange in a secure manner.

25 In the preceding description, various aspects of the present invention have been described. For purposes of explanation, specific numbers, systems and configurations were set forth in order to provide a thorough understanding of the present invention. However, it is apparent to one skilled in the art having the benefit of this disclosure that the present invention may be practiced without the 30 specific details. In other instances, well-known features were omitted or simplified in order not to obscure the present invention.

Embodiments of the present invention may be implemented in hardware or software, or a combination of both. However, embodiments of the invention may be implemented as computer programs executing on programmable systems comprising at least one processor, a data storage system (including 5 volatile and non-volatile memory and/or storage elements), at least one input device, and at least one output device. Program code may be applied to input data to perform the functions described herein and generate output information. The output information may be applied to one or more output devices, in known fashion. For purposes of this application, a processing system embodying the 10 portions of the present invention includes any system that has a processor, such as, for example, a digital signal processor (DSP), a microcontroller, an application specific integrated circuit (ASIC), or a microprocessor.

The programs may be implemented in a high level procedural or object oriented programming language to communicate with a processing system. The 15 programs may also be implemented in assembly or machine language, if desired. In fact, the invention is not limited in scope to any particular programming language. In any case, the language may be a compiled or interpreted language.

The programs may be stored on a removable storage media or device 20 (e.g., floppy disk drive, read only memory (ROM), CD-ROM device, flash memory device, digital versatile disk (DVD), or other storage device) readable by a general or special purpose programmable processing system, for configuring and operating the processing system when the storage media or device is read by the processing system to perform the procedures described herein. 25 Embodiments of the invention may also be considered to be implemented as a machine-readable storage medium, configured for use with a processing system, where the storage medium so configured causes the processing system to operate in a specific and predefined manner to perform the functions described herein.

30 Although the operations describe herein may be described as a sequential process, some of the operations may in fact be performed in parallel or

concurrently. In addition, in some embodiments the order of the operations may be rearranged without departing from the spirit of the invention.

The techniques described herein are not limited to any particular hardware or software configuration; they may find applicability in any computing or processing environment. The techniques may be implemented in hardware, software, or a combination of the two. The techniques may be implemented in programs executing on programmable machines such as mobile or stationary computers, personal digital assistants, set top boxes, cellular telephones and pagers, and other electronic devices, that each include a processor, a storage medium readable by the processor (including volatile and non-volatile memory and/or storage elements), at least one input device, and one or more output devices. Program code is applied to the data entered using the input device to perform the functions described and to generate output information. The output information may be applied to one or more output devices. One of ordinary skill in the art may appreciate that the invention can be practiced with various computer system configurations, including multiprocessor systems, minicomputers, mainframe computers, and the like. The invention can also be practiced in distributed computing environments where tasks may be performed by remote processing devices that are linked through a communications network.

Each program may be implemented in a high level procedural or object oriented programming language to communicate with a processing system. However, programs may be implemented in assembly or machine language, if desired. In any case, the language may be compiled or interpreted.

Program instructions may be used to cause a general-purpose or special-purpose processing system that is programmed with the instructions to perform the operations described herein. Alternatively, the operations may be performed by specific hardware components that contain hardwired logic for performing the operations, or by any combination of programmed computer components and custom hardware components. The methods described herein may be provided as a computer program product that may include a machine readable medium having stored thereon instructions that may be used to program a processing

system or other electronic device to perform the methods. The term "machine readable medium" used herein shall include any medium that is capable of storing or encoding a sequence of instructions for execution by the machine and that cause the machine to perform any one of the methods described herein.

5 The term "machine readable medium" shall accordingly include, but not be limited to, solid-state memories, optical and magnetic disks, and a carrier wave that encodes a data signal. Furthermore, it is common in the art to speak of software, in one form or another (e.g., program, procedure, process, application, module, logic, and so on) as taking an action or causing a result. Such
10 expressions are merely a shorthand way of stating the execution of the software by a processing system cause the processor to perform an action of produce a result.

While this invention has been described with reference to illustrative embodiments, this description is not intended to be construed in a limiting sense.

15 Various modifications of the illustrative embodiments, as well as other embodiments of the invention, which are apparent to persons skilled in the art to which the invention pertains are deemed to lie within the spirit and scope of the invention.